ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Complexity 21 (2005) 487–501

Journal of COMPLEXITY

www.elsevier.com/locate/jco

Symplectic methods for the approximation of the exponential map and the Newton iteration on Riemannian submanifolds

Jean-Pierre Dedieu*, Dmitry Nowicki

MIP. Département de Mathématique, Université Paul Sabatier, 31062 Toulouse Cedex 04, France

Received 20 February 2004; received in revised form 23 June 2004; accepted 6 September 2004
Available online 11 February 2005

Abstract

We use a Hamiltonian approach and symplectic methods to compute the geodesics on a Riemannian submanifold.

Keywords: Submanifold; Geodesic; Exponential map; Newton iteration; Symplectic; Hamiltonian

1. Introduction

Let V be a p-dimensional Riemannian real complete manifold. In this paper we study computational aspects of the Newton method for finding zeros of smooth mappings $f : V \to \mathbb{R}^p$. The Newton operator is defined by

$$N_f(x) = \exp_x(-Df(x)^{-1} f(x)). \tag{1}$$

Here $\exp_x : T_x V \to V$ is the exponential map, which "projects" the tangent space at x on the manifold. The Newton method has two important properties: fixed points for N_f correspond to zeros for f and the convergence of the Newton sequences ($x_0 = x$ and $x_{k+1} = N_f(x_k)$) is quadratic for any starting point x in a neighborhood of a nonsingular zero.

* Corresponding author.
E-mail addresses: dedieu@mip.ups-tlse.fr (J.-P. Dedieu), nowicki@mip.ups-tlse.fr (D. Nowicki).

0885-064X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2004.09.010

When $V = \mathbb{R}^n$, the exponential map is just a translation: $\exp_x(u) = x + u$ and the Newton operator has the usual form:

$$N_f(x) = x - Df(x)^{-1} f(x)$$

but for a general manifold this is no more true. Except for some cases the exponential map has no analytic expression and we have to compute it numerically: this is the main subject of this paper.

Newton method for maps or vector fields defined on manifolds has already been considered by many authors: Shub [29] defines Newton's method for the problem of finding the zeros of a vector field on a manifold and uses retractions to send a neighborhood of the origin in the tangent space onto the manifold itself. Udriste [35] studies Newton's method to find the zeros of a gradient vector field defined on a Riemannian manifold; Owren and Welfert [27] define Newton iteration for solving the equation $f(x) = 0$ where f is a map from a Lie group to its corresponding Lie algebra; Smith [34] and Edelman et al. [10] develop Newton and conjugate gradient algorithms on the Grassmann and Stiefel manifolds. Shub [30], Shub and Smale [31–33], see also, Blum et al. [4], Malajovich [22], Dedieu and Shub [7] introduce and study the Newton method on projective spaces and their products. Another paper on this subject is Adler et al. [3] where qualitative aspects of Newton method on Riemannian manifolds are investigated for both mappings and vector fields. This paper contains an application to a geometric model for the human spine represented as a 18-tuple of 3×3 orthogonal matrices. Recently, Ferreira–Svaiter [11] give a Kantorovich like theorem for Newton method for vector fields defined on Riemannian manifolds and Dedieu et al. [6] study alpha-theory for both mappings and vector fields.

The computation of the exponential map depends mainly on the considered data structure. In some cases the exponential is given explicitely (Euclidean or projective spaces, spheres …) or may be computed via linear algebra packages (the orthogonal group, Stiefel or Grassmann manifolds [10,34]). The classical description uses local coordinates and the second order system which gives the geodesic curve $x(t)$ with initial conditions $x(0) = x$, and $\dot{x}(0) = u$:

$$\ddot{x}_i(t) + \sum_{j,k} \Gamma^i_{jk} \dot{x}_j(t) \dot{x}_k(t) = 0, \quad 1 \leqslant i \leqslant n,$$

$$x(0) = x, \quad \dot{x}(0) = u.$$

In these equations Γ^i_{jk} are the Christoffel symbols and the exponential is equal to $\exp_x(u) = x(1)$, see Do Carmo [9] or others textbooks on this subject: Dieudonné [8], Gallot et al. [13], Helgason [17]. Such an approach is used by Noakes [25] who considers the problem of finding geodesics joining two given points. We notice that the computation of local coordinates and of the Christoffel symbols may be itself a very serious problem and depends again on the data structure giving the manifold V.

In [5] Celledoni and Iserles consider the approximation of the exponential for finite dimentional Lie groups contained in the general linear group using splitting techniques. Munthe et al. [36] approximate the matrix exponential by the use of a generalized polar decomposition. See also Munthe et al. [36] for the generalized polar decomposition on Lie groups, Krogstad et al. [21] and Iserles et al. [19].

In this paper we concentrate our efforts on submanifolds. Let $F : U \to \mathbb{R}^m$ be a C^2 map where $U \subset \mathbb{R}^n$ is open. Let V denote its zero set: $V = F^{-1}(0) \subset U \subset \mathbb{R}^n$. We suppose that $DF(x) : \mathbb{R}^n \to \mathbb{R}^m$ is onto for each $x \in U$. In that case, V is a C^2 submanifold contained in $\mathbb{R}^n$ and its dimension is equal to $p = n - m$. V is equipped with the Riemannian structure inherited from $\mathbb{R}^n$: the scalar product on $T_x V$ is the restriction of the usual scalar product in $\mathbb{R}^n$. This case is particularly important in optimization theory when V, the set of feasible points, is defined by equality constraints. In this framework, to compute the geodesic curves with initial value conditions, we take a mechanical approach: a geodesic is the trajectory of a free particle attached to the submanifold V, see Abraham–Marsden [1] or Marsden–Ratiu [23]. We give a first description of this trajectory in terms of Lagrangian equations and then, via an optimal control approach and Pontryagin's maximum principle, in terms of Hamiltonian equations. Our numerical methods are based on this last system: we use symplectic methods to solve it (second-, fourth- or sixth-order Gauss method).

We are now able to compute the Newton operator attached to a system of equations defined on V, say $f : V \to \mathbb{R}^{n-m}$. The last section is devoted to numerical examples. We compare this Riemannian Newton method (called here GNI for "Geometric Newton Iteration") with the usual Euclidean Newton method (called CNI for "Classical Newton Iteration") which solves the extended system $f(x) = 0$, and $F(x) = 0$ with $x \in \mathbb{R}^n$. Both methods, for these examples, give comparable results with a smaller number of iterates for the GNI and a slightly better accuracy for the CNI.

Other numerical methods for problems posed on Riemannian manifolds require the computation of the exponential map. This will be the purpose of a second paper. We thank here Luca Amodei for valuable discussions about this symplectic approach.

2. The equations defining the geodesics

The exponential map $\exp_x : T_x V \to V$ is defined in the following way: for $x \in V$ and $u \in T_x V$ let $x(t), t \in \mathbb{R}$, be the geodesic curve such that $x(0) = x$ and $\dot{x}(0) = \frac{dx(t)}{dt}|_{t=0} = u$. Then $\exp_x(u) = x(1)$. Let us denote by $N_x V$ the normal space at x. We have

$$T_x V = \operatorname{Ker} DF(x) \quad \text{and} \quad N_x V = (T_x V)^{\perp} = \operatorname{Im} DF(x)^*.$$

This geodesic is characterized by the following system:

$$\begin{aligned} &x(t) \in V, \\ &\ddot{x}(t) \in N_{x(t)} V, \\ &x(0) = x,\ \dot{x}(0) = u. \end{aligned} \tag{2}$$

We introduce a Lagrange multiplier $\lambda(t) \in \mathbb{R}^m$ so that system (2) becomes

$$\begin{aligned} &F(x(t)) = 0, \\ &\ddot{x}(t) = -DF(x(t))^* \lambda(t), \\ &x(0) = x,\ \dot{x}(0) = u. \end{aligned} \tag{3}$$

This geodesic curve may be interpreted as the trajectory of a free particle attached to V. Using the formalism of Lagrangian mechanics, see Marsden–Ratiu [23, Section 8.3], we notice

that this system is given by the Euler–Lagrange equation associated with the following Lagrangian:

$$\mathcal{L}(x, \dot{x}, \lambda) = \frac{1}{2}\|\dot{x}\|^2 - \sum_{i=1}^{m} \lambda_i F_i(x) \tag{4}$$

that is

$$\begin{aligned} &F(x(t)) = 0, \\ &\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial \mathcal{L}}{\partial x}, \\ &x(0) = x,\ \dot{x}(0) = u. \end{aligned} \tag{5}$$

Definition 2.1. For a linear operator $A : E \to F$ between two Euclidean spaces, we denote by $A^\dagger$ its generalized inverse. It is the composition of three maps, $A^\dagger = i \circ B^{-1} \circ \Pi_{\operatorname{Im} A}$ with $\Pi_{\operatorname{Im} A}$ the orthogonal projection from F onto $\operatorname{Im} A$, $B : (\operatorname{Ker} A)^\perp \to \operatorname{Im} A$ the restriction of A, $i : (\operatorname{Ker} A)^\perp \to E$ the canonical injection.

The operator $AA^\dagger$ is equal to the orthogonal projection $F \to \operatorname{Im} A$ and $A^\dagger A$ is the orthogonal projection $E \to (\operatorname{Ker} A)^\perp$. When A is onto one has $A^\dagger = A^*(AA^*)^{-1}$ and $AA^\dagger = \mathrm{id}_F$ while, when A is injective, $A^\dagger = (A^*A)^{-1}A^*$ and $A^\dagger A = \mathrm{id}_E$.

Proposition 2.1. *For any $x \in V$ and $u \in T_x V$ system (3) is equivalent to*

$$\begin{aligned} &\ddot{x}(t) = -DF(x(t))^\dagger D^2F(x(t))(\dot{x}(t), \dot{x}(t)), \\ &\lambda(t) = \left(DF(x(t))DF(x(t))^*\right)^{-1} D^2F(x(t))(\dot{x}(t), \dot{x}(t)), \\ &x(0) = x,\ \dot{x}(0) = u. \end{aligned} \tag{6}$$

Proof. To obtain (6) from (3) we differentiate two times $F(x(t)) = 0$ so that

$$D^2F(x(t))(\dot{x}(t), \dot{x}(t)) + DF(x(t))\ddot{x}(t) = 0.$$

By (3) we get

$$D^2F(x(t))(\dot{x}(t), \dot{x}(t)) - DF(x(t))DF(x(t))^*\lambda(t) = 0.$$

Since $DF(x(t))$ is onto, $DF(x(t))DF(x(t))^*$ is nonsingular and this gives $\lambda(t)$ and $\ddot{x}(t)$. Conversely, (6) gives

$$\ddot{x}(t) = -DF(x(t))^\dagger D^2F(x(t))(\dot{x}(t), \dot{x}(t)) = -DF(x(t))^*\lambda(t).$$

Moreover

$$DF(x(t))\ddot{x}(t) = -D^2F(x(t))(\dot{x}(t), \dot{x}(t))$$

that is

$$\frac{d^2}{dt^2}F(x(t)) = 0.$$

This gives

$$F(x(t)) = F(x(0)) + DF(x(0))\dot{x}(0) + \frac{1}{2}\int_0^t \frac{d^2}{dt^2}F(x(s))ds$$

$$= F(x) + DF(x)u + 0 = 0. \qquad \square$$

Let us now introduce the Hamilton equations. To obtain them we consider the problem of finding a minimizing geodesic with two given endpoints as the following optimal control problem (see [35]):

$$\min \int_0^T \|y(t)\|^2 dt$$

subject to the constraints $\dot{x}(t) = y(t)$, $x(0) = x_0$, $x(1) = x_1$, $F(x(t)) = 0$ for every $t \in [0, T]$, where x_0 and x_1 are given points in V. According to Pontryagin's maximum principle, the Hamiltonian for problems like

$$\min \int_0^T f_0(x, y, t)dt$$

subject to the constraints $\dot{x} = f(x, y, t)$, $x(0) = x_0$, $x(1) = x_1$, $F(x(t)) = 0$ for every $t \in [0, T]$, can be written as

$$\mathcal{H}(p, x, \mu) = -f_0 + \langle p, f\rangle + \sum_{i=1}^m \mu_i DF_i(x)\dot{x}.$$

In our case we obtain

$$\mathcal{H}(x, p, \mu) = \langle p, \dot{x}\rangle - \frac{1}{2}\|\dot{x}\|^2 + \sum_{i=1}^m \mu_i DF_i(x)\dot{x}$$

with $p \in \mathbb{R}^n$, $\mu \in \mathbb{R}^m$. The Hamilton equations are

$$\dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial x}(x(t), p(t), \mu(t)),$$

$$p(t) = \dot{x}(t) - DF(x(t))^*\mu(t),$$

$$\dot{\mu}(t) = -\lambda(t), \quad \mu(0) = 0. \tag{7}$$

Proposition 2.2. *Let $x \in V$ and $u \in T_xV$ be given. System (7) is equivalent to*

$$\dot{p}(t) = -\sum_{i=1}^m \mu_i D^2F_i(x(t))\dot{x}(t),$$

$$\dot{x}(t) = \Pi_{T_{x(t)}V}p(t) = \Big(id - DF(x(t))^\dagger DF(x(t))\Big)\, p(t),$$

$$\mu(t) = -DF(x(t))^{*\dagger} p(t),$$

$$x(0) = x, \quad p(0) = u \tag{8}$$

which is also equivalent to system (3).

Proof. To obtain (8.1) from (7) we differentiate $\mathcal{H}$ with respect to x to obtain

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial x} &= \left\langle p, \frac{\partial \dot{x}}{\partial x} \right\rangle - \left\langle \dot{x}, \frac{\partial \dot{x}}{\partial x} \right\rangle + \sum_{i=1}^{m} \mu_i DF_i(x) \frac{\partial \dot{x}}{\partial x} + \sum_{i=1}^{m} \mu_i D^2 F_i(x)\dot{x} \\ &= \sum_{i=1}^{m} \mu_i D^2 F_i(x)\dot{x}\end{aligned}$$

by (7.2). The two other equations in (8) are obtained from (7.2) by projecting $p(t)$ on $\operatorname{Ker} DF(x(t))$ and $\operatorname{Ker} DF(x(t))^{\perp} = \operatorname{Im} DF(x(t))^*$ so that

$$\dot{x}(t) = \Pi_{T_{x(t)V}} p(t)$$

and

$$-DF(x(t))^* \mu(t) = \Pi_{\operatorname{Im} DF(x(t))^*} p(t).$$

Since $DF(x(t))$ is injective we get

$$\begin{aligned}\mu(t) &= DF(x(t))^{*\dagger} DF(x(t))^* \mu(t) \\ &= -DF(x(t))^{*\dagger} \Pi_{\operatorname{Im} DF(x(t))^*} p(t) = -DF(x(t))^{*\dagger} p(t).\end{aligned}$$

To obtain (8.1) from (3) and (6) we differentiate (7) to obtain

$$\begin{aligned}\dot{p}(t) &= \ddot{x}(t) - DF(x(t))^* \dot{\mu}(t) - \sum_{i=1}^{m} \mu_i D^2 F_i(x(t))\dot{x}(t) \\ &= \ddot{x}(t) + DF(x(t))^* \lambda(t) - \sum_{i=1}^{m} \mu_i D^2 F_i(x(t))\dot{x}(t) = -\sum_{i=1}^{m} \mu_i D^2 F_i(x(t))\dot{x}(t).\end{aligned}$$

The initial condition (8.4) is given by

$$p(0) = \dot{x}(0) - DF(x(0))^* \mu(0) = u.$$

To obtain (3) from (8) we differentiate $\dot{x}(t) = p(t) + DF(x(t))^* \mu(t)$ to get

$$\ddot{x}(t) = \dot{p}(t) + \sum_{i=1}^{m} \mu_i D^2 F_i(x(t))\dot{x}(t) + DF(x(t))^* \dot{\mu}(t) = DF(x(t))^* \lambda(t).$$

By the same equation we get

$$\dot{x}(0) = p(0) + DF(x(0))^* \mu(0) = u.$$

Moreover,

$$DF(x(t))\dot{x}(t) = DF(x(t))\Pi_{\operatorname{Ker} DF(x(t))} p(t) = 0,$$

by integrating we get

$$F(x(t)) = F(x(0)) + \int_0^t DF(x(s))\dot{x}(s)ds = 0$$

and we are done. □

3. Numerical integration of Hamilton equations

3.1. Symplectic Runge–Kutta methods

To integrate the Hamiltonian system (8) we use symplectic Runge–Kutta methods. We do not use partitioned Runge–Kutta methods like Stormer–Verlet because our Hamiltonnian is not separable. Let us consider an autonomous system:

$$\dot{y} = G(y) : U \to \mathbb{R}^p \tag{9}$$

defined over an open set $U \subset \mathbb{R}^p$. In the case considered here $y = (x, p) \in \mathbb{R}^n \times \mathbb{R}^n$ and $G(x, p)$ is given by (8).

Let us denote by $\Phi_t(y)$ the associated integral flow: $y(t) = \Phi_t(y)$ is the solution of (9) with the initial condition $y(0) = y$. The implicit Runge–Kutta method we have implemented is given by

$$\begin{aligned} &y_0 = y(0),\\ &y_{k+1} = y_k + \tau \sum_{i=1}^{s} b_i G(Y_i),\\ &Y_i = y_k + \tau \sum_{j=1}^{s} a_{ij} G(Y_j),\ 1 \leqslant i \leqslant s. \end{aligned} \tag{10}$$

Here $\tau > 0$ is the given step size, s is a given integer, $Y_1, \ldots, Y_s$ are auxilliary variables, $(a_{ij})_{1 \leqslant i,j \leqslant s}$, $(b_i)_{1 \leqslant i \leqslant s}$ are the coefficients defining the considered method. For our experiments we use Gauss methods of order 2, 4 and 6. The corresponding coefficients are

- Order 2: $s = 1$, $a_{11} = \frac{1}{2}$ and $b_1 = 1$.
- Order 4: $s = 2$,

a_{ij}		
	$\frac{1}{4}$	$\frac{1}{4} - \frac{\sqrt{3}}{6}$
	$\frac{1}{4} + \frac{\sqrt{3}}{6}$	$\frac{1}{4}$
b_i	$\frac{1}{2}$	$\frac{1}{2}$

- Order 6: $s = 3$,

a_{ij}			
	$\frac{5}{36}$	$\frac{2}{9} - \frac{\sqrt{15}}{15}$	$\frac{5}{36} - \frac{\sqrt{15}}{30}$
	$\frac{5}{36} + \frac{\sqrt{15}}{24}$	$\frac{2}{9}$	$\frac{5}{36} - \frac{\sqrt{15}}{24}$
	$\frac{5}{36} + \frac{\sqrt{15}}{30}$	$\frac{2}{9} + \frac{\sqrt{15}}{15}$	$\frac{5}{36}$
b_i	$\frac{5}{18}$	$\frac{4}{9}$	$\frac{5}{18}$

see Hairer et al. [16] or Sanz et al. [28] about these methods.

Let us denote $\psi_t : \mathbb{R}^p \to \mathbb{R}^p$ which outputs y_{k+1} in terms of y_k. Let us consider again the case of (8) with $y = (x, p)$. The properties of these methods are the following:

- They are symplectic, i.e.

$$\omega^2(\psi_t(y)) = \omega^2(y)$$

for any Hamiltonian system and for any $\tau > 0$ where ω^2 is the differential 2-form

$$\omega^2 = \sum_{i=1}^{n} dx_i \wedge dp_i. \tag{11}$$

To solve Eqs. (10) we have chosen a successive approximation scheme. These iterations are convergent when the following inequality:

$$\|DG\|\|A\|\tau < 1$$

is satisfied. The norm of DG could be estimated by a direct derivation of the right-hand side of system (7). This lead to the following inequality:

$$\|DG\| \leqslant \|D^2F\|^2\|(DFDF^*)^{-1}\| + \|D^2F\|^2\|(DFDF^*)^{-1}\|^2\|DF\|^2 + \|DF\|\|D^3F\|\|(DFDF^*)^{-1}\|.$$

Let v denotes the index of internal iteration inside the kth step. For a given tolerance *tol* our termination criterion is

$$\|Y^{v+1} - Y^v\| \leqslant \tau.tol.$$

In our experiments we have chosen $tol \approx 10^{-8}$. Further decreasing of the tolerance did not lead to better accuracy of the geometric Newton method.

3.2. Backward error analysis

We apply the backward error analysis techniques from Hairer–Lubich [15] and Hairer [14] to the case of system (8) integrated by symplectic methods. We show that the computed points x_k ($y_k = (x_k, p_k)$) are arbitrarily close to the geodesic corresponding to a nearby

Riemannian structure and the same initial conditions as in the exact problem. We also estimate the distance between these two Riemannian distances. More precisely

Theorem 3.1. *Let $V \subset \mathbb{R}^n$ be a Riemannian submanifold defined by the equation $F(x) = 0$, where $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is analytic on a certain neighbourhood of D in $\mathbb{C}^n$. Let I be a symplectic numerical integrator of order r with a sufficiently small step size $\tau > 0$. Let us denote by $g(x)$ the $n \times n$ positive definite matrix defining the Riemannian metric at $x \in V$. Then V can be endowed with a new Riemannian metric $\tilde{g}(x, \tau)$ such that*

1. $$\|g(x) - \tilde{g}(x, \tau)\| = O(\tau^{r+1}). \tag{12}$$
2. *There exists $\tau^* > 0$ such that for any initial condition*

$$x(0) = x_0 \in V, \ \dot{x}(0) = u_0 \in T_{x_0}V \tag{13}$$

we have

$$\|x_k - \tilde{x}(k\tau)\| = O\left(\exp\left(-\frac{\tau^*}{2\tau}\right)\right) \tag{14}$$

for any k such that $k\tau \leqslant T = \tau exp\,(\tau^*/2\tau)$, where x_k is the numerical solution provided by the integrator and $\tilde{x}(t)$ is the exact geodesic associated with the metric $\tilde{g}(x, \tau)$ and the same initial conditions (13).

Proof. We apply the Corollary 6 from Hairer–Lubich [15] to our system. See also Hairer [14] where constrained Hamiltonian systems are considered. From this corollary, we get a Hamiltonian

$$\tilde{H}(x, p, \tau) = H(x, p) + O(\tau^{r+1}) \tag{15}$$

such that its trajectories satisfy estimate (14). The Riemannian metric $\tilde{g}$ is built from the kinetic energy of this new Hamiltonian (see [2, Chapter 9]). This metric satisfies inequality (12). The trajectories of the Hamiltonian 15 are the geodesics of this metric. □

4. The Newton operator

How do we compute a Newton step? Let us first recall the geometric context.

Let $F : U \to \mathbb{R}^m$ be a C^2 map where $U \subset \mathbb{R}^n$ is open. Let V denote its zero set: $V = F^{-1}(0) \subset U \subset \mathbb{R}^n$ and let $f : U \to \mathbb{R}^{n-m}$ be given, as smooth as necessary. We also denote by $\hat{f}$ its restriction to V.

To compute the Newton operator

$$N_{\hat{f}}(x) = \exp_x(-D\hat{f}(x)^{-1}\hat{f}(x))$$

we need the derivative $D\hat{f}(x) : T_xV \to \mathbb{R}^{n-m}$. This derivative is the projection onto the tangent space T_xV of the derivative $Df(x) : \mathbb{R}^n \to \mathbb{R}^{n-m}$. Since this projection is equal to $I - DF(x)^\dagger DF(x)$ (see Definition 2.1) we obtain

$$D\hat{f}(x) = \pi_{T_xV} Df(x) = (I - DF(x)^\dagger DF(x))Df(x).$$

5. Experimental results

Example 5.1. Quadratic manifold.

In this example we consider the quadratic manifold:

$$V = \left\{ x \in \mathbb{R}^{100} \ : \sum_{k=1}^{5} \frac{x_k^2}{k} - \sum_{k=6}^{100} \frac{x_k^2}{k} = 1 \right\}.$$

To compute the geodesics we use the Gauss method of order 4 with $\tau = 0.01$. On this manifold we solve the following problems:

1. A linear system $Bx = 0$, where B is a random 99×100 matrix,
2. A quadratic system: $Bx + c\|x\|^2 = 0$, where c is a given random vector in $\mathbb{R}^{99}$.

The initial point $x_0 \in \mathbb{R}^{100}$ of the Newton sequence is taken at random in following sense. Each component $x_{0k}, k = 1 \ldots 99$, is taken randomly in $[-1, 1]$ with respect to the uniform distribution and $x_{0,100}$ is computed to satisfy the equation defining V.

The corresponding results are displayed in Table 1 in the column "Geometric Newton Iteration" or "GNI".

For comparison we also display the results obtained for the same problems using the classical Newton method to the extended system

$$(F, f) : U \to \mathbb{R}^m \times \mathbb{R}^{n-m}.$$

We call "Classical Newton Iteration" or "CNI" the corresponding sequence.

The typical behaviour of these iterations is shown in Fig. 1 and Table 1. They show a quadratic convergence obtained in few steps. Then we reach a limit due to round-off errors. The number of Newton steps is better for the geometric method than for the classical one but the precision is better for the classical method than for the geometric one. This is due to the amount of computation which is more important for the GNI.

Example 5.2. "Distorted" quadratic manifolds.

In this example we consider the following manifold:

$$V = \left\{ x \in \mathbb{R}^{100} \ : \sum_{k=1}^{5} \frac{\sin(x_k)^2}{k} - \sum_{k=6}^{100} \frac{\sin(x_k)^2}{k} = 1 \right\}.$$

Table 1
Results for the quadratic manifold

Problem	# of GNI steps	GNI Precision	# of CNI steps	CNI Precision
Linear	7	1.01×10^{-14}	8	2.2×10^{-16}
Quadratic	10	1.06×10^{-15}	12	3.0×10^{-16}

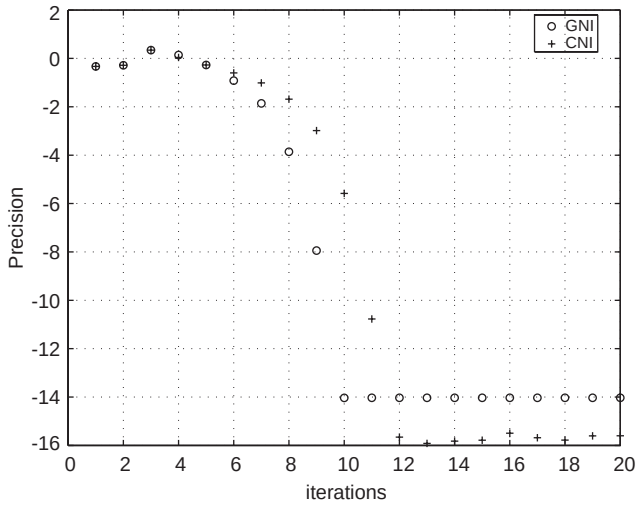

Fig. 1. Convergence of geometric and classical Newton methods for the quadratic manifold. ∘ stands for GNI and + for CNI.

Table 2
Results for the distorted quadratic manifold

Problem	# of GNI steps	GNI Precision	# of CNI steps	CNI Precision
Linear	7	7.1×10^{-15}	11	1.4×10^{-16}
Quadratic	9	2.1×10^{-14}	12	2.2×10^{-16}

On this manifold we solved the same test problems with the same parameters as in Example 1. This manifold has an infinite number of connected components. We restrict our study to the connected component V_0 such that $|x_i| \leqslant \pi/2$ for each $x \in V_0$. The initial point was taken randomly in V_0 like in the previous example. Under these conditions two solutions were found. The corresponding results are displayed in Table 2.

Example 5.3. Katsura's system.

The following equations appear in a problem of magnetism in physics. For more details see Katsura–Sasaki [20], and also the web site [37].

$$\begin{aligned} &u_m = \sum_{i=-N}^{N} u_i u_{m-i} \ ; \ m = 0 \ldots N-1, \\ &\sum_{i=-N}^{N} u_i = 1, \\ &u_{-m} = u_m \ ; \ m = 1 \ldots 2N-1, \\ &u_m = 0; m = N+1 \ldots 2N-1. \end{aligned} \tag{16}$$

Table 3
Results for Katsura's example

M	# of GNI steps	Precision
0 (CNI)	14	1.78×10^{-16}
1	12	4.4×10^{-14}
5	9	6.7×10^{-15}
10	9	5.6×10^{-15}
20	7	1.0×10^{-14}
40	7	3.2×10^{-15}

After eliminating u_m for $m \notin 0 \dots N$ we obtain $N+1$ equations in $\mathbb{R}^{N+1}$:

$$u_m = \sum_{i=m+1}^{N} u_i u_{i-m} + \sum_{i=0}^{N-m} u_i u_{i+m} + \sum_{i=1}^{m} u_i u_{m-i} \; ; \; m = 0 \dots N-1,$$
$$2\sum_{i=1}^{N} u_i + u_0 = 1. \tag{17}$$

This system is not a priori posed on a manifold. For this reason we split the equations into two groups: the M first equations from (17) (for $m = 0 \dots M-1$) define a manifold V_M of codimension M, and the remaining equations are considered as a system on V_M. The GNI starts at a random point $x_0 \in V_M$. To find such a point we take at random a point y_0 in a box containing V_M (such a box is easy to compute from the structure of Katsura's system). Then we "project" y_0 on V via the Newton–Gauss method in $\mathbb{R}^{N+1}$.

In Table 3 we display the results for $N = 40$ and different values for M. We use the 4-order Gauss numerical integrator with $\tau = 0.01$. The results for the classical Newton method are also included: they correspond to the codimension $M = 0$. We do not know the number of real solutions of this system. During the test we found four different solutions.

In Fig. 2, we illustrate the same example with $N = 2$ and $M = 1$. The four first GNI iterates are located on the surface: $x_0, x_{1g}, x_{2g}, x_{3g}$ and x_{4g} while the iterates corresponding to the CNI ($x_0, x_{1c}, x_{2c}, x_{3c}, x_{4c}$) are clearly located outside the surface. We notice the same facts as in our first example: better numerical behaviour for the CNI but a better complexity in terms of the number of iterates for the GNI.

Example 5.4. The generalized Brown system.

This example is a generalizaion of Brown's system, see Floudas–Pardolos [12, Chapter 14, test problem 14.1.5]:

$$\sum_{\substack{1 \leqslant i \leqslant N, \\ i \neq k}} x_i + 2x_k^r = N+1, \quad k = 1 \dots N-1$$
$$\prod_{1 \leqslant i \leqslant N} x_i = 1. \tag{18}$$

The case $N = 5$ and $r = 1$ corresponds to the original system.

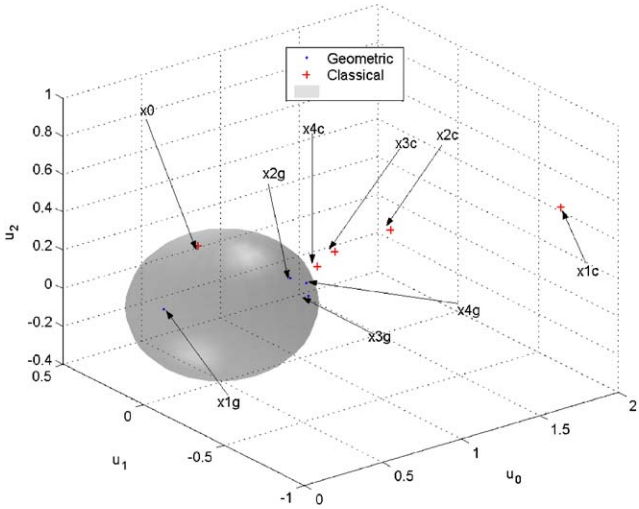

Fig. 2. Convergence of CNI and GNI for the Katsura's example, $N = 2,\ M = 1$.

Table 4
Results for generalized Brown example

M	# of GNI steps	Precision
0 (CNI)	8	2.3×10^{-15}
1	7	4.7×10^{-14}
3	6	7.2×10^{-14}
5	6	7.6×10^{-14}
9	6	1.8×10^{-14}

Like in the previous example, this system is not a priori posed on a manifold. Its equations are also split into two groups in the same way.

The GNI starts at a random point x_0 in the manifold V of codimension M. To find such a point we take at random a point y_0 in the box $0.5 \leqslant x_k \leqslant 1.5$, $k = 1 \ldots N$ with respect to the uniform distribution. Then, we project y_0 on V via the Newton–Gauss method in $\mathbb{R}^N$.

In Table 4 we display the results for $N = 10$, $r = 3$, and different values for M. We use the 4-order Gauss numerical integrator with $\tau = 0.01$. The results for the classical Newton method are also included: they correspond to the codimension $M = 0$. We do not know the number of real solution of this system. During the test we found the solution: $x = (1, \ldots, 1)^T$.

References

[1] R. Abraham, J. Marsden, Foundations of Mechanics, Addison-Wesley, Reading, MA, 1978.
[2] V.I. Arnold, Mathematical Methods of Classical Mechanics, Springer, New York, 1978.

[3] R. Adler, J.-P. Dedieu, J. Margulies, M. Martens, M. Shub, Newton method on Riemannian manifolds and a geometric model for the human spine, IMA J. Numer. Anal. 22 (2002) 1–32.
[4] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer, Berlin, 1998.
[5] E. Celledoni, A. Iserles, Methods for the approximation of the matrix exponential in a Lie-algebraic setting, IMA J. Numer. Anal. 21 (2001) 463–488.
[6] J.-P. Dedieu, G. Malajovich, P. Priouret, Newton method on Riemannian manifolds: covariant alpha-theory, IMA J. Numer. Anal. 23 (2003) 395–419.
[7] J.-P. Dedieu, M. Shub, Multihomogeneous Newton's method, Math. Comput. 69 (2000) 1071–1098.
[8] J. Dieudonné, Treatise on Analysis, Academic Press, New York, 1988.
[9] M. Do Carmo, Riemannian Geometry, Birkhauser, Boston, 1992.
[10] A. Edelman, T. Arias, S. Smith, The Geometry of Algorithms with Orthogonality Constraints, SIAM J. Matrix Anal. Appl. 20 (1998) 303–353.
[11] O. Ferreira, B. Svaiter, Kantorovich's theorem on Newton's method in Riemannian manifolds, J. Complexity 18 (2002) 304–329.
[12] C.A. Floudas, P. Pardolos, et al., Handbook of Test Problems in Local and Global Optimization, Kluwer Academic Publishers, Dordrecht, 1999.
[13] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry, Springer, Berlin, 1993.
[14] E. Hairer, Global modified Hamiltonian for constrained symplectic integrators, Numer. Math. 95 (2) (2003) 325–336.
[15] E. Hairer, C. Lubich, The life-span of backward error analysis for numerical integrators, Numer. Math. 77 (2) (1997) 325–336.
[16] E. Hairer, S.P. Norsett, G. Wanner, Solving ordinary differential equations. I. Nonstiff problems, Springer Series in Computational Mathematics, vol. 8, second ed., Springer, Berlin, 1993.
[17] K. Helgason, Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, New York, 1979.
[19] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørset, A. Zanna, Lie-group methods, Acta Numer. 9 (2001) 215–365.
[20] S. Katsura, M. Sasaki, The asymmetric continuous distribution function of the effective field of the Ising model in the spin glass and the ferromagnetic states on the Bethe lattice, Physica A 157-3 (1989) 1195–1202.
[21] S. Krogstad, H. Munthe-Kaas, A. Zanna, Generalized Polar Coordinates on Lie Groups and Numerical Integrators, Technical Report No. 244, Department of Informatics, University of Bergen, 2003. Numer. Math., to appear.
[22] G. Malajovich, On generalized Newton algorithms, Theoret. Comput. Sci. 133 (1994) 65–84.
[23] Marsden, J., Ratiu, T., Introduction to Mechanics and Symmetry, Teubnor-Archiv Zur Mathematic, vol. 17, Springer, Berlin, 1994.
[24] H.Z. Munthe-Kaas, G.R.W. Quispel, A. Zanna, Generalized polar decompositions on Lie groups with involutive automorphisms, Found. Comput. Math. 1 (2001) 297–324.
[25] L. Noakes, A Global Algorithm for Geodesics, J. Austral. Math. Soc. Ser. A 64 (1998) 37–50.
[27] B. Owren, B. Welfert, The Newton iteration on Lie groups, BIT 40 (1) (2000) 121–145.
[28] J.M. Sanz-Serna, M.P. Calvo, Numerical Hamiltonian Problems, Chapman & Hall, New York, 1994.
[29] M. Shub, Some remarks on dynamical systems and numerical analysis, in: L. Lara-Carrero, J. Lewowicz (Eds.), Dynamical Systems and Partial Differential Equations, Proceedings of VII ELAM Equinoccio, Universidad Simon Bolivar, Caracas, 1986, pp. 69–92.
[30] M. Shub, Some remarks on Bezout's theorem and complexity, in: J.M. Marsden, M.W. Hirsch, M. Shub (Eds.), From Topology to Computation: Proceedings of the Smalefest, Springer, Berlin, 1993, pp. 443–455.
[31] M. Shub, S. Smale, Complexity of Bézout's theorem I: geometric aspects, J. Am. Math. Soc. 6 (1993) 459–501.
[32] M. Shub, S. Smale, Complexity of Bezout's theorem IV: probability of success, extensions, SIAM J. Numer. Anal. 33 (1996) 128–148.
[33] M. Shub, S. Smale, Complexity of Bézout's theorem V: polynomial time, Theoret. Comput. Sci. 133 (1994) 141–164.

[34] S. Smith, Optimization Techniques on Riemannian Manifolds, Fields Institute Communications, vol. 3, American Mathematical Society, Providence, RI, 1994, pp. 113–146.
[35] C. Udriste, Convex Functions and Optimization Methods on Riemannian Manifolds, Kluwer, Dordrecht, 1994.
[36] A. Zanna, H.Z. Munthe-Kaas, Generalized polar decompositions for the approximation of the matrix exponential, Siam J. Matrix Anal. Appl. 23 (2001) 840–862.
[37] http://www-sop.inria.fr/saga/POL/BASE/2.multipol/katsura.html.